

CERN-TH.7020/93  
 UND-HEP-93-BIG 02  
 September 1993

## Weak Annihilation and the End Point Spectrum in Semileptonic B Decays

I.I.Bigi

*Theoretical Physics Division, CERN  
 CH-1211 Geneva 23, Switzerland*

and

*Dept.of Physics, University of Notre Dame du Lac, Notre Dame, IN 46556, U.S.A.* <sup>1</sup>

*e-mail address: VXCERN::IBIGI, BIGI@UNDHEP*

N.G.Uraltsev

*Dept.of Physics, University of Notre Dame du Lac, Notre Dame, IN 46556*

and

*St.Petersburg Nuclear Physics Institute, Gatchina, St.Petersburg 188350, Russia* <sup>1</sup>

*e-mail address: URALTSEV@UNDHEP, URALTSEV@LNPI.SPB.SU*

### Abstract

A general relationship is formulated between the contributions of ‘Weak Annihilation’ (WA) to inclusive semileptonic decays of heavy flavour hadrons and the matrix elements of four fermion operators. We argue that nonperturbative contributions from low energy hadronic final states provide the dominant impact of WA on the semileptonic  $b \rightarrow u$  width of  $B^-$  mesons. In that case WA affects the lepton spectrum mainly in its endpoint region – i.e. beyond the kinematical boundary for  $b \rightarrow c$  transitions – and we expect the later to look quite different in  $B^-$  than in  $B_d$  decays. At present we are unable to make a truly quantitative prediction; yet a detailed experimental comparison of the  $B^-$  and the  $B_d$  lepton spectra will enable one to determine the size of the relevant matrix elements and to check directly to what extent the factorization approximation works here. Using universal analyticity properties of the decay amplitude we analyze WA in the perturbative regime and rederive our earlier results about the absence of a power enhancement due to gluon emission in a more general way. We address also the problem of the invariant mass of hadrons in the final state in semileptonic  $B$  decays and identify nontrivial cancellations in  $\langle M_{\text{had}}^2 \rangle$ .

---

<sup>1</sup>permanent address

# 1 Introduction

In previous papers [1] we have shown how the inclusive decays of heavy flavour hadrons can be treated reliably in QCD via an expansion in  $1/m_Q$ ,  $m_Q$  being the mass of the heavy flavour quark. The leading term in this expansion represents – not surprisingly – the results of the simple spectator ansatz. We have found that there are corrections of order  $1/m_Q$  to the *shape* of spectra, but not to totally integrated rates. Those receive corrections first on the  $1/m_Q^2$  level and they are of an intriguing nature: they distinguish between baryon decays on one hand and meson decays on the other while affecting the latter in a way that is practically independent of the flavour of the spectator antiquark. On the  $1/m_Q^3$  level the conventional spectator-dependent effects arise, namely ‘Weak Annihilation’ (WA), ‘Pauli Interference’ (PI) and ‘Weak Scattering’ (WS) in baryons<sup>1</sup>; these represent preasymptotic corrections to a parton model treatment whose size depends on the flavour of the ‘spectator’ quarks or antiquarks in a given heavy flavour hadron  $H_Q$ .

It is the purpose of this note to study in more detail the role played by such spectator dependent effects; we will analyze almost exclusively WA with the emphasis on nonperturbative effects that may eliminate chirality suppression and thus enhance its effect to an appreciable level. The discussion will center mainly on semileptonic decays since those allow a more reliable and more detailed theoretical treatment. Among other things one can go beyond total rates and study lepton *spectra*. An analysis of their observed shape will, at least in principle, allow to extract the size of hadronic matrix elements that in turn determine the impact of WA on other quantities.

The paper will be organized as follows: after a general overview of the effects associated with WA we set the framework for our theoretical discussion in Section 2 where we establish the connection between the effects of WA and matrix elements of four fermion operators that appear in quark diagrams on the tree level. Such a relationship was implicitly assumed in earlier papers [2, 3]; its existence was later explicitly conjectured and supported by the perturbative analysis of ref. [4]. Here we present a general analysis focussing on nonperturbative effects. It was shown in ref. [4] that the power enhancements produced by gluons cancel each other in *inclusive* rates. We illustrate the general formalism adopted here by using it to rederive these results; we demonstrate that they are indeed – as stated in ref. [4] – of rather general nature. In Section 3 we apply these findings to the semileptonic spectra and show that the interesting effects originate in the endpoint region and occur mainly there. Still there are ‘tail’ contributions coming from the endpoint effects; in fact they reflect the ‘hybrid’ renormalization of the operators that was first discussed for inclusive effects in ref. [3]. In Section 4 we treat the invariant mass distribution of the hadronic final state before summarizing our study in Section 5. We will discuss primarily beauty

---

<sup>1</sup> Quite often a distinction is made between the exchange of a  $W$  boson in the  $s$  and in the  $t$  channel referred to as Weak Annihilation and  $W$  exchange, respectively. We find such a classification somewhat artificial because both interactions lead to very similar four fermion couplings; they may differ in their colour flow at *tree* level, yet this distinction fades away under gluon renormalization. For that reason we have adopted the classification proposed in ref. [2]: WA denotes then the exchange of  $W$  bosons both in the  $s$  and  $t$  channel of a quark-antiquark system whereas WS refers to  $W$  exchange driving quark-quark scattering in baryons.

decays with some comments added on charm decays. It is worth noting that the first, though rather general, study of the analyticity properties of the amplitudes for semileptonic heavy flavour decays has been done in ref. [5]; some details omitted in the present paper can be found there.

WA must be understood in a more general way [4] than the one suggested by simple parton level diagrams. A difference in the lifetimes of  $B_d$  and  $B^-$  mesons can be produced by PI as well as by WA. PI requires the presence of at least two identical quarks or antiquarks. WA on the other hand is associated with the presence of a quark-antiquark pair of the same flavour; the antiquark is originally present in the  $B$  meson as a constituent and the quark gets created at the weak interaction vertex. In semileptonic  $B$  decays PI cannot intervene and the whole difference in  $\Gamma_{SL}(B^-)$  vs.  $\Gamma_{SL}(B^0)$ , according to the existing classification, has to come from WA via the KM suppressed  $b \rightarrow u$  transitions.

The simplest annihilation process is shown in Fig.1a and is easily calculated; we will refer to it as ‘parton annihilation’. It scales like  $(\mu_{\text{hadr}}/m_b)^3$  with the mass of the heavy quark. Yet rather subtle dynamical effects have to be addressed once one goes beyond the simplest diagram. For example  $\bar{q}q$  pairs can annihilate into a gluon, Fig.1b; the emergence of such new final states in the decay represents a conventional annihilation effect. However this is not the only modification – for annihilation affects at the same time the production rates for the ‘old’ states as illustrated in Figs.1c: the original ‘quasifree’ amplitude now interferes with the annihilation-induced amplitude. This modification does not represent the square of any particular amplitude and therefore can actually be negative thus *decreasing* the width. Nevertheless this latter kind of effect has to be included under the term WA unless one wants to coin a new name for such spectator-sensitive corrections!

The amplitudes for a final state with a gluon or with a  $\bar{q}q$  pair when taken *separately* are enhanced by inverse powers of light quark masses (or more generally energies  $\langle E_q \rangle$ ) as compared to the parton estimate of WA. This enhancement [8, 3, 4] can be understood as being caused by final state interactions, namely the mixing between nearly mass degenerate flavour singlet  $\bar{q}q$  pairs and gluons<sup>2</sup>. It is quite natural to expect that this enhancement evaporates when one sums over these decay processes. It was shown in ref. [4] by explicit calculation that this is indeed the case. It means by the way that the two possible understandings of WA are dramatically different when the decaying quark is heavy enough.

Of course the physical states coming from flavour singlet  $\bar{q}q$  pairs in practice cannot be distinguished from those generated purely by gluons; i.e. observable transitions cannot exhibit an  $1/\langle E_q \rangle$  enhancement. The decays of the  $B_c$  meson consisting of  $b\bar{c}$  offer an intriguing exception to this general rule: for final states containing a  $c$  and  $\bar{c}$  pair can be distinguished from those without heavy quarks and the conventional WA reaction  $B_c \rightarrow d\bar{u}g$  can thus be isolated. On the other hand the overall impact of annihilation in our more general formulation is hard to extract: for one cannot any longer rely on a comparison of  $\tau(B_c)$  vs.  $\tau(B_u)$  as an unambiguous yardstick since due to  $m_c \gg m_u$  no (light) flavour symmetry can be invoked. There is however a saving grace: both constituents of the  $B_c$  meson are heavy and therefore we know

---

<sup>2</sup>N.U. is grateful to J.Collins for the discussion of this point.

how to explicitly compute the relevant transition rates. The situation in  $B_s$  decays falls somewhere in between the decays of  $B_c$  and  $B_{u,d}$  mesons.  $SU(3)$  breaking effects in total beauty decay widths are under control here [6, 7]; yet at the same time a reliable isolation of  $s\bar{s}$  states produced (or affected) by WA is questionable if possible at all in a real experiment.

The analysis of ref. [4] actually yielded a threefold result concerning WA:

- All gluonic enhancements of the form  $(1/\langle E_q \rangle)^n$  that appear in individual diagrams cancel out in the inclusive transition rates.
- The emission of gluons carrying momenta larger than the typical hadronic scale  $\mu_{\text{hadr}}$  provides the ‘hybrid’ renormalization [3] of the local four fermion operators that drive the spectator-dependent effects in the decays.
- The chirality suppression still persists for WA when treating gluon radiation perturbatively in the leading *log* approximation.

The last observation implies that the weight of WA is reduced in heavy flavour meson decays, but not necessarily to insignificant levels. For helicity suppression could be overcome or its impact at least be softened by  $\mathcal{O}(\alpha_s)$  subleading contributions from the emission of hard gluons, or via nonperturbative effects. Numerically significant effects can be expected from the latter, and we will discuss them now in detail for the case of semileptonic  $b \rightarrow ul\nu$  transitions.

## 2 General Framework; Total Rates

This Section consists of four main parts: (A) We will state a relationship between the contributions of WA to inclusive rates and the matrix elements of four fermion operators as it arises in the approach formulated in refs. [2, 3]. (B) We will discuss the contributions to WA from low mass hadronic final states. (C) We will analyze how higher energy states contribute to WA and show that their impact is given by the corresponding perturbative corrections to the matrix elements of these operators. (D) Some consequences for nonleptonic decays will be pointed out.

(A) Following the notation of ref. [5] where the traditional approach to deep inelastic scattering was modified for the case of semileptonic heavy flavour decays, one starts from the tensor

$$R_{\mu\nu}(q) = \int d^4z \exp(-iqz) \cdot \langle B | iT(J_\mu^+(z)J_\nu(0)) | B \rangle = -\delta_{\mu\nu}R_1 + \\ + p_\mu p_\nu R_2 + i\epsilon_{\mu\nu\rho\lambda} p_\rho q_\lambda R_3 + q_\mu q_\nu R_4 + (p_\mu q_\nu + q_\mu p_\nu)R_5 + i(p_\mu q_\nu - q_\mu p_\nu)R_6 \quad ; \quad (1)$$

$q$  and  $p$  are the momenta carried by the lepton pair  $l\nu$  and the  $B$  meson, respectively;  $J_\mu$  denotes the underlying quark current. (One has  $J_\mu = \bar{b}\gamma_\mu(1 - \gamma_5)u$  when studying WA in semileptonic  $B$  decays.) The scalar quantities  $R_i$  are functions of the Lorentz invariants  $q^2$  and  $x = 1 - (p \cdot q)/M_B^2$ . In the  $B$  restframe  $(1 - x)M_B$  is the energy of the lepton pair whereas  $xM_B$  represents the energy of the final hadronic state<sup>3</sup>. To compute decay rates one contracts the absorptive part of  $R_{\mu\nu}$  with the corresponding leptonic tensor  $L_{\mu\nu}$  built from the leptonic charged currents. The absorptive part

---

<sup>3</sup>For convenience we have adopted a different definition for the energy variable than used in ref. [5].

of  $R_{\mu\nu}$ , denoted by  $W_{\mu\nu}$ , is given by the discontinuities of the functions  $R_i$  in the complex variable  $x$  in the physical region. For example the total semileptonic width is given by

$$\Gamma_{\text{sl}} \propto \int_0^1 dq^2 \int_{\frac{1-q^2/M_B^2}{2}}^{1-\sqrt{q^2/M_B^2}} dx \sqrt{(1-x)^2 M_B^2 - q^2} \cdot \tilde{W}(x, q^2) ,$$

$$\tilde{W}(x, q^2) = q^2 W_1(x, q^2) + \frac{M_B^2}{3} (M_B^2 (1-x)^2 - q^2) W_2(x, q^2) = \frac{1}{3} (q_\mu q_\nu - q^2 \delta_{\mu\nu}) W_{\mu\nu}(q) \quad (2)$$

where Eq.(2) has been written for the case of massless leptons<sup>4</sup>.

To establish a relationship between the correlator function of eq.(1) and the local product of the corresponding quark currents, we start from a general expression for the time-ordered product of any two operators:

$$\begin{aligned} & \int d^4 z \exp(-iqz) \cdot \langle B | iT(O_1(z)O_2(0)) | B \rangle = \\ & = \sum_n \frac{\langle B | O_1(z_0=0) | n \rangle \cdot \langle n | O_2(0) | B \rangle}{-E_B + E_n + q_0 - i\epsilon} + \sum_{n'} \frac{\langle B | O_2(0) | n' \rangle \cdot \langle n' | O_1(z_0=0) | B \rangle}{-E_B + E_{n'} - q_0 - i\epsilon} , \quad (3) \end{aligned}$$

where  $n, n'$  are intermediate states with energies  $E_{n,n'}$  and  $E_B$  denotes the energy of the  $B$  meson. Integrating the imaginary part of eq.(3) over the energy  $q_0$  from  $E_B - E_2$  to  $E_B - E_1$  yields

$$\begin{aligned} & \int_{E_B-E_2}^{E_B-E_1} dE \frac{1}{\pi} \text{Im} \langle B | iT(O_1 O_2) | B \rangle_{E, \vec{q}} = \\ & = \sum_{E_1 < E_n < E_2} (2\pi)^3 \delta^3(\vec{p}_n - \vec{p}_B + \vec{q}) \langle B | O_1(z_0=0) | n \rangle \cdot \langle n | O_2(0) | B \rangle + \\ & + \sum_{2E_B-E_2 < E_{n'} < 2E_B-E_1} (2\pi)^3 \delta^3(\vec{p}_{n'} - \vec{p}_B - \vec{q}) \langle B | O_2(0) | n' \rangle \cdot \langle n' | O_1(z_0=0) | B \rangle . \quad (4) \end{aligned}$$

Since

$$\begin{aligned} & \langle B | J_\mu^+(z_0=0) J_\nu(z_0=0) | B \rangle_{\vec{q}} = \\ & = (2\pi)^3 \sum_n \delta^3(\vec{p}_n - \vec{p}_B + \vec{q}) \langle B | J_\mu^+(z_0=0) | n \rangle \cdot \langle n | J_\nu(z_0=0) | B \rangle \quad (5) \end{aligned}$$

one arrives at the following conclusion: integrating the discontinuity of the Greens function in eq.(1) over some finite energy interval yields the contribution to the matrix element of the *equal time* product of the hadron currents from intermediate states having kinematically allowed energies for this energy interval. This relationship holds as long as  $u$  channel singularities (at  $x > 1$ ) in the time ordered Greens functions do not contribute in the energy interval considered. Integrating also over the space momenta  $\vec{q}$  one finally obtains a completely local product of currents.

After this general remark we return to the semileptonic width as given by eq.(2). The various intermediate states saturating the correlators  $W$  are characterized by

---

<sup>4</sup>Note the explicit nonanalytical factor  $\sqrt{(1-x)^2 M_B^2 - q^2} = |\vec{q}|$  that did not appear in the discussion of ref. [5].

rather different energy and momentum scales; the dynamics associated with them is therefore quite different and we will discuss them separately. For the sake of definiteness we shall assume in what follows that we are in the rest frame of the  $B$  meson (or one where its motion is at most nonrelativistic). We anticipate the major contribution to WA to come from nonperturbative effects, and to show up in low lying states. So we shall start with discussing the range of low  $E_n \sim \mu_{\text{hadr}}$ , i.e.  $x \sim \mu_{\text{hadr}}/m_b \ll 1$ .

**(B)** When both the energy and the momentum of the hadronic final state are small compared to  $M_B$  the lepton pair in the decay carries a large invariant mass,  $q^2 \simeq m_b^2$ , making the leptonic part of the amplitude (including the  $l\nu$  phase space factor) basically insensitive to both the energy and the momentum of the state  $|n\rangle$ . Its value can thus be evaluated at  $q_\mu \simeq p_\mu = (M_B, 0, 0, 0)$ ; summing over low lying states with a constant weight factor given by the leptonic tensor  $L_{\mu\nu}(p)$  one actually performs an integration of the functions  $\tilde{W}$  over both energy and momentum. The hadronic part of the decay probability is thus indeed determined by the matrix element of the local operator  $\bar{u}_L \gamma_\mu b_L(0) \cdot \bar{b}_L \gamma_\nu u_L(0)$ . We then draw the following conclusion: *The width for semileptonic decays into hadronic final states with energies substantially smaller than  $M_B$  is described by the contributions of these states to the matrix element of the local product of the underlying weak quark currents, taken between  $B$  meson states –  $\langle B | \bar{u}_L \gamma_\mu b_L(0) \cdot \bar{b}_L \gamma_\nu u_L(0) | B \rangle$  – and multiplied by the leptonic tensor evaluated at  $q^2 = M_B^2$ .* Corrections due to the small (for  $E_n \ll m_b$ ) variation in the lepton pair momentum can be taken into account by expanding the leptonic part of the amplitude around  $q^2 = M_B^2$ . They are described by similar four fermion operators containing derivatives; their impact is suppressed by extra powers of  $\mu_{\text{hadr}}/m_b$ .

Since our focus is on width differences due to WA we consider the difference of the matrix elements taken between charged and neutral mesons:

$$M_{\mu\nu} = 4(\langle B^- | \bar{u}_L \gamma_\mu b_L \cdot \bar{b}_L \gamma_\nu u_L | B^- \rangle - \langle \bar{B}^0 | \bar{u}_L \gamma_\mu b_L \cdot \bar{b}_L \gamma_\nu u_L | \bar{B}^0 \rangle) = f_B^2 (v p_{B_\mu} p_{B_\nu} - g M_B^2 \delta_{\mu\nu}) \quad (6)$$

i.e. the size of WA is expressed in terms of two dimensionless parameters  $v$  and  $g$ .

The contributions of WA to the width are obtained, as already stated, by contracting  $M_{\mu\nu}$  with the leptonic tensor  $L_{\mu\nu}(p_B)$  given by the absorptive part of the  $l\nu$  loop

$$L_{\mu\nu}(q) = -\frac{1}{6\pi} (1 - m_l^2/q^2) \left[ \left(1 - \frac{m_l^2}{2q^2} - \frac{m_l^4}{2q^4}\right) q^2 \delta_{\mu\nu} - \left(1 + \frac{m_l^2}{q^2} - \frac{2m_l^4}{q^4}\right) q_\mu q_\nu \right] \quad (7)$$

where  $m_l$  denotes the charged lepton mass; the correlator of the leptonic charged currents  $L_{\mu\nu}$  of course has a transverse structure for  $m_l = 0$ . For the contribution of WA to semileptonic  $B$  decays one then finds

$$\Delta\Gamma_{\text{SL}} = \Gamma_{SL}(B^-) - \Gamma_{SL}(B^0) \simeq \frac{G_F^2 |V_{ub}|^2 f_B^2 M_B^3}{8\pi} \left(1 - \frac{m_l^2}{M_B^2}\right) \left(v \frac{m_l^2}{M_B^2} + 2g\right) \quad (8)$$

The part of the hadronic matrix element which is proportional to  $p_\mu p_\nu$  yields a result that vanishes in the limit of  $m_l \rightarrow 0$ : this reflects the conventional chirality suppression of WA as it applies to semileptonic  $B$  decays. On the other hand the part in  $M_{\mu\nu}$

proportional to  $\delta_{\mu\nu}$  does not suffer from such a reduction; its weight then depends on the size of  $g$ . In general  $g \neq 0$  and therefore the impact of WA on semileptonic decays  $b \rightarrow u$  could be significant as seen from the following comparison:

$$\frac{\Gamma_{SL}(B^-) - \Gamma_{SL}(B^0)}{\Gamma(B \rightarrow \tau\nu_\tau)} \simeq \frac{2gM_B^2}{m_\tau^2} \sim 6 \quad \text{for } g \sim \frac{1}{3} . \quad (9)$$

The simplest annihilation decay is  $B \rightarrow l\nu_l$ ; it is represented by  $|n\rangle = |0\rangle$  in the current correlator decomposition of eq.(3) and correspondingly by the vacuum factorization ansatz for the matrix element  $M_{\mu\nu}$ , eq.(6), of the local four fermion operator. The contribution of the decay  $B \rightarrow l\nu_l$  to  $\Gamma(B^-) - \Gamma(B^0)$  is therefore given exactly by the factorizable piece of the underlying matrix element, i.e. in the notation of eq.(6) by

$$v_{\text{vac}} = 1 \quad , \quad g_{\text{vac}} = 0 . \quad (10)$$

Helicity suppression applies and the dominant mode in this category is  $B \rightarrow \tau\nu$ . Yet in general  $g \neq 0$  will hold for the matrix element in eq.(6) and helicity suppression can thus be expected to be vitiated when hadrons are present in the final state; the relevant question is on which *numerical* level will this arise.

(C) Up to this point we have discussed low energy hadronic final states where a perturbative treatment is never applicable; their combined effect is to generate a (nonperturbative) value for the quantities  $v$  and  $g$ . When higher momenta arise in the hadronic final state one can consider analyzing the structure of inclusive hadronic states and estimate their contributions by using perturbative QCD that operates in terms of quark and gluon fields. Such an analysis was first undertaken in a similar context in ref. [8] with the focus mainly on overcoming helicity suppression. Only the one-gluon intermediate state was considered there: its amplitude was found to be free of helicity suppression and in addition to be enhanced by a factor of  $m_b^2/\langle E_{sp} \rangle^2$  relative to the WA width without gluons, with  $\langle E_{sp} \rangle$  being an average energy of the light (anti)quark. Yet in ref. [4] we have shown that such an enhancement disappears from the *fully inclusive* amplitude, i.e. the latter is *regular* in the limit of vanishing spectator quark momenta. We actually found that for  $\mu_{\text{hadr}} \ll E \ll m_b$  (with  $E$  denoting the energy of the hadronic states) the WA contribution is of the form  $\alpha_s dE/E$ . Upon integration over this energy range it contributes to the so-called ‘hybrid’ renormalization of the underlying four fermion operator. In other words, summing over all hadronic final states leads to the renormalization of the operator entering eq.(6) down to the low scale  $\sim \mu_{\text{hadr}}^2$ .

An explicit calculation has shown – somewhat surprisingly – that the hybrid renormalization does *not* modify the Lorentz structure of the four quark operator. Therefore perturbative corrections *per se* do not induce a  $g$  term for the matrix element in eq.(6) in the leading log approximation (LLA) and helicity suppression is thus found to persist on that level! It can be overcome only at a relatively small level  $\mathcal{O}(\alpha_s(m_b^2)/\pi)$  beyond LLA – unless it emerges via nonperturbative dynamics as was discussed above.

It is instructive to see schematically how the cancellation of the  $1/\langle E_{sp} \rangle^2$  and  $1/\langle E_{sp} \rangle$  terms occurs and the stated structure of the perturbative corrections emerges by using rather general considerations which are based on the analytical properties of the correlator in eq.(1). It will demonstrate – as was stated in ref. [4] – that the

underlying argument is more general than the simple quark diagram that is encountered in the explicit calculation of the perturbative corrections; it will be of use for our subsequent discussion of semileptonic spectra. For the sake of simplicity we will ignore here the colour quantum numbers that strictly speaking forbid these gluonic effects to first order in  $\alpha_s$  for semileptonic decays; in fact the diagrams we consider are literally those that were treated in ref. [4] in addressing nonleptonic decays.

The first nontrivial perturbative correction to the annihilation width is obtained for example from the diagram in Fig.2; we consider it in greater detail because it could *a priori* lead to the most singular corrections. The width given by this diagram is most conveniently obtained by keeping the value of  $q^2$  at first fixed and integrate over  $x$ , i.e. over the energy; subsequently one performs the integration over  $q^2$  according to eq.(2). This diagram generates the correlator function  $R$  with the following obvious analytical properties in the complex  $x$  plane: there are three poles corresponding to the massless gluon and to on shell  $\bar{u}u$  states with  $\bar{u}$  being the spectator and  $u$  coming from the  $b \rightarrow u$  transition. These  $\bar{u}u$  pairs have an invariant mass squared of  $(k + p_{\text{sp}})^2 \simeq 2(k \cdot p_{\text{sp}})$ . This is much smaller than  $m_b^2$  even for a high momentum  $|\vec{k}| \sim m_b$ . The three poles are therefore very closely spaced in the  $x$  plane:

$$x_1 \simeq \frac{1 - Q^2}{2} \quad , \quad x_{2,3} \simeq \frac{1 - Q^2}{2} + \frac{(p_u \cdot p_{\text{sp}})}{m_b^2} \quad (11)$$

as indicated in Fig.3; we have introduced here the dimensionless ratio  $Q^2 = q^2/m_b^2$ . Integrating the imaginary part of this diagram (which in this particular case amounts to summing the three  $\delta$ -functions) can be done by evaluating the integral over path  $\alpha$  in Fig.3. The standard procedure is then to move the integration contour into the complex plane away from the singularities, like path  $\beta$ . As long as that contour does not approach the singular points  $x$  near  $(1 - Q^2)/2$  one can safely set the momenta of the spectator quark to zero – for it merely shifts the exact position of the poles by a small amount. This order  $\alpha_s$  contribution to the total width therefore possesses a regular limit for vanishing spectator momenta; it can be calculated by integrating the function containing the triple pole at the point  $x \simeq (1 - Q^2)/2$  over a contour like  $\beta$ . Accordingly the decay amplitude cannot contain any inverse powers of the spectator mass or energy [4]. The explicit evaluation of this diagram is thus in principle straightforward; yet in practice it requires attention to technicalities like the numerators in the fermion propagators that restore the correct dimensionality for this diagram. They have the generic structure  $(p_b - q)_\alpha(p_b - q)_\beta$ ; therefore they either cancel one of the three poles  $1/(1 - Q^2 - 2x)$  in the amplitude or instead provide a factor  $\vec{q}^2 = ((1 - x)^2 - Q^2)m_b^2$  that will vanish at the point  $\sqrt{Q^2} = 1 - x$  where the Jacobian in eq.(2) is singular. An explicit expression is easily obtained by stretching the integration contour around the cut produced by the Jacobian  $\sqrt{(1 - x)^2 - Q^2}$  in eq.(2) as represented by the contour  $\gamma$  in Fig.3; alternatively one can use the relation

$$\int dx f(x) \frac{1}{\pi} \text{Im} \frac{1}{(x - x_o + i\epsilon)^{n+1}} = -\frac{1}{n!} f^{(n)}(x_o) \quad .$$

Upon differentiating the Jacobian  $\sqrt{(1 - x)^2 - Q^2}$  one obtains for both expressions of the numerators in the propagators a result  $\propto \frac{1}{1 - Q^2} = M_B/2|\vec{k}|$  where  $\vec{k}$  is seen



as the momentum of the massless quark in the semileptonic decay at a given value of the lepton invariant mass  $q^2$ . The expression indeed becomes singular when the singularity of the function pinches the cut of the Jacobian corresponding to the edge of phase space, i.e. at the maximum invariant mass of the lepton pair. The integration over  $q^2$  in this region produces the hybrid renormalization  $\log(m_b^2/\mu^2)$ .

So far we have discussed WA in terms of the annihilation of free quarks. Yet one needs the matrix element of the current correlator between  $B$  hadrons, eq.(1). Let us consider the matrix element of the one gluon annihilation transition between a real  $B$  meson. Due to the distribution in the relative motion of the spectator (anti)quark and the heavy quark inside the hadron the three original poles in the amplitude get smeared out and actually turn into cuts. Higher order corrections are expected to produce a similar effect. Yet the general arguments proceed as before as long as the invariant mass of the intermediate state is small compared to  $m_b$  (or more exactly to the energy released into the final hadronic states). For this ensures that the singularities are shifted from their original position  $x_o = (1 - Q^2)/2$  by a much smaller amount than its distance to the Jacobian singularity  $x = 1 - \sqrt{Q^2}$  which reflects the edge of phase space. The smallness of the mass of the intermediate state is the only essential fact one has to invoke from a quark model description of the heavy hadron; in the naive constituent quark model this mass squared indeed cannot exceed the scale  $\mu_{\text{hadr}} \cdot m_b$ . As long as such an inequality holds and the momenta of the final hadronic states are much smaller than  $m_b$  one can deduce as before that the resulting contributions to the decay width of the hadron are given by the corresponding corrections to the matrix element of the local product of currents. This is the reason why the perturbative effects of the one gluon annihilation process are reproduced correctly by calculating the corresponding *Euclidean* graph for the *local* four fermion operator [4].

The situation would appear to change somewhat when higher order corrections are considered; in particular the pole corresponding to the gluon propagator changes drastically and the integral over the invariant mass of the gluon fragmentation will extend over a wider range. None of this will however spoil the cancellation of the enhanced terms  $\propto m_b^2/\langle E_{sp} \rangle^2$  and  $m_b/\langle E_{sp} \rangle$ . For the amplitude computed via the contour  $\beta$  in Fig.3 is determined by momenta that are high compared to  $\mu_{\text{hadr}}$ ; the overall effect is therefore determined by the ultraviolet properties of the parton diagram rather than by details of the real hadronic states (as long as the invariant mass of the lepton pair does not approach the mass of  $b$  quark too closely).

**(D)** The previous discussion provides the formal framework for the QCD treatment of WA in semileptonic decays. A similar analysis – although of course not so rigorous – can be applied to nonleptonic heavy flavour decays as well with obvious technical modifications: one starts from the correlator function generated by the four fermion Lagrangians rather than the weak currents in eq.(1). To obtain total widths one considers the correlator at  $q = 0$  and applies an operator expansion in  $1/m_b$  to it. The leading spectator dependent effects are obtained when one treats a quark-antiquark (for WA) or a quark-quark (for PI or WS) pair in the intermediate state as free fields. The general structure thus obtained resembles very much that found in semileptonic decays; in particular the analogous four fermion operators appear

then, albeit with a somewhat more general Lorentz and colour structure. Subleading corrections – both perturbative and nonperturbative ones – to nonleptonic and semileptonic rates do differ and *a priori* are expected to be larger in the former than in the latter. Yet it is important that they reflect the same underlying structure; their relative weight both for perturbative and nonperturbative corrections is controlled by the same quantity, namely the distance from the thresholds of the free parton model graph.

Our preceeding discussion has yielded two important phenomenological conclusions: **(i)** The actual size of WA for semileptonic  $b \rightarrow u$  decays is indeed determined by the matrix element  $\langle B^- | \bar{u}_L \gamma_\mu b \cdot \bar{b} \gamma_\nu u | B^- \rangle$ ; nonleptonic decays are governed by similar matrix elements with different colour contracting scheme<sup>5</sup>. **(ii)** Helicity suppression persists in the presence of *perturbative* gluon corrections, even when ‘hybrid’ renormalization is included in the leading log approximation.

On the other hand helicity suppression can be expected to be vitiated by subleading perturbative or by nonperturbative corrections. This is to say in the notation of eq.(6) that in general a term proportional to  $\delta_{\mu\nu}$  will be induced, albeit with a  $g$  presumably somewhat smaller than unity. Numerically sizeable WA contributions to inclusive rates can presumably be expected only from nonperturbative dynamics. WA could then still change the width of semileptonic  $b \rightarrow u$  decays in  $B^-$  by up to twenty percent relative to that in  $B^0$ . Such an effect would not be of purely academic interest due to the following observation: one would expect – if nonperturbative effects are the main driving force overcoming helicity suppression – that WA would populate semileptonic final states with mainly low energy hadronic systems. WA would thus contribute the bulk of its weight in the relatively narrow endpoint region of semileptonic decays which in turn would magnify its impact there. In the next section we describe the impact of WA on semileptonic *spectra* in more details.

### 3 Lepton spectra

The energy spectrum for the charged leptons in semileptonic beauty decays can be expressed in terms of the correlation function given in eq.(1):

$$\begin{aligned} \frac{d\Gamma}{dy} &= \frac{G_F^2 |V_{ub}|^2}{32\pi^4} M_B^5 \cdot \int_0^y dQ^2 \int_{\frac{1-Q^2}{2}}^{\frac{1-Q^2}{2} + \frac{1}{2}(1-y)(1-Q^2/y)} dx \tilde{V}(x, Q^2) = \\ &= \frac{G_F^2 |V_{ub}|^2}{32\pi^4} M_B^5 \cdot \int_{(1-y)/2}^{1-y/2} dx \int_0^{2y(1-x-y/2)} dQ^2 \tilde{V}(x, Q^2) , \\ \tilde{V}(x, Q^2) &= \frac{q^2}{M_B^2} W_1(x, q^2) + (y(1-x - \frac{y}{2}) M_B^2 - \frac{1}{2} q^2) W_2(x, q^2) + (1-x-y) q^2 W_3(x, q^2) \end{aligned} \quad (12)$$

with  $y = 2E_l/M_B$  |<sub>c.m. frame</sub> denoting the normalized charged lepton energy; lepton masses have been ignored here. The lepton spectrum in charm decays is described by the same formula with two obvious modifications:  $M_B$  is replaced by  $M_D$ , and

---

<sup>5</sup>This suggestion was formulated in ref. [4] although it had been implicitly implied already in analysis of ref. [3].

the positive sign in front of the  $W_3$  term in eq.(12) is changed into a negative one reflecting the  $V + A$  structure of the charge conjugated lepton current. Integrating eq.(12) over  $y$  reproduces of course eq.(2) for the total semileptonic width.

Different kinematical regions of this energy spectrum are shaped by different dynamical regimes controlling the recoiling hadronic system. As before we will first consider low energy states affected by nonperturbative dynamics. Low mass hadronic states whose energy does not exceed some small value  $M_{\text{np}}$  can contribute only to a rather limited slice in the lepton energy:  $1 - 2M_{\text{np}}/M_B \lesssim y < 1$ ; this interval becomes more and more narrow as the mass of the heavy quark increases and actually shrinks to a point in the limit of an infinitely heavy quark. A simple minded valence quark description suggests that  $M_{\text{np}} \sim m_{\text{sp}} \sim \mu_{\text{hadr}}$ , i.e. this interval of the lepton energy remains constant in absolute units. On the other hand the main effect of WA as analyzed in the previous section is to occur in this small region! The contribution from WA then appears in the lepton energy distribution effectively like a  $\delta$ -function near  $y = 1$  with a smearing in the energy of order  $\mu_{\text{hadr}}/m_b$ . (It is worth noting that similar effects have been identified in ref. [9] for flavour-independent preasymptotic corrections.) This contribution to the total width is thus concentrated in the region A in Fig.4 and its overall strength is determined by the value of  $g$  in the matrix element eq.(6). The *shape* of the spectrum inside this region on the other hand depends on the details of hadronization into the low-lying hadronic states. At present neither of these can be calculated reliably from QCD.

If the energy released into the hadronic system is large compared to the typical invariant mass of the latter then one can rely on a perturbative treatment. The quark model suggests for the square of this mass as a typical value  $M_{\text{annih}}^2 \simeq |\vec{q}| \cdot m_{\text{sp}}$ ; the aforementioned requirement is thus indeed satisfied even for large three-momenta. Accordingly we may rely on a perturbative treatment of WA below region A. (There are actually two subregions shown in Fig.4 – B and C – which are separated by the point  $y = 1 - M_D^2/M_B^2$ ;  $b \rightarrow c$  transitions are restricted to region C.) The structure of the perturbative amplitude can again be simplified by considering it in the limit of vanishing spectator momenta; the functional dependence of the result can then be readily obtained from the discussion of the previous section. For example the contribution to  $\tilde{V}$  coming from the part of the amplitude containing a double pole is given by  $\delta'(\frac{1-Q^2}{2} - x)$ , and the integration over  $x$  in eq.(12) yields merely  $\delta((1-y)(1-Q^2/y))$ ; the integration over  $Q^2$  then leads to

$$\frac{d\Gamma}{dy} \simeq c \frac{\alpha_s}{1-y} \quad \text{for} \quad 1-y > M_{\text{annih}}^2/m_b^2 \sim \mu_{\text{hadr}}/m_b, \quad (13)$$

where the factor  $c$  comes from the “ $\delta$ -function” term in the spectrum region A. Eq.(13) gives the hybrid logarithm  $\log m_b^2/\mu_{\text{hadr}}^2$  when integrated over the whole spectrum and in fact reflects the nonfactorizable contribution to the hybrid renormalization of the operator in eq.(6). Factorizable corrections describing (virtual) corrections to individual semileptonic vertices also produce such terms. Their effects however do not depend on the lepton energy  $y$ ; they lead to the known renormalization of both weak currents in the matrix element eq.(6) down to the low scale  $\mu_{\text{hadr}}^2$ . These in principle increase the height of the spectrum in region A by a factor  $\alpha^4$  (given explicitly below

in eq.(14)) over the naive quark model result. This has been incorporated in eq.(6) via the explicit factor  $f_B^2$  undergoing the same corrections [3]. The WA contribution over the whole spectrum is sensitive to both types of hybrid renormalization.

After this general overview of the structure of the perturbative corrections we give explicit expressions. Four fermion operators as they appear in eq.(6) do not change their Lorentz structure upon hybrid renormalization [3, 4] as long as they are constructed from purely left-handed fields; yet they mix with the colour octet counterparts:

$$\begin{aligned} \bar{u}_L \gamma_\mu b_L \cdot \bar{b}_L \gamma_\nu u_L &\rightarrow (\mathfrak{a}^{9/2} - 1/9(\mathfrak{a}^{9/2} - 1)) \cdot (\bar{u}_L \gamma_\mu b_L)(\bar{b}_L \gamma_\nu u_L) - \\ &- 2/3(\mathfrak{a}^{9/2} - 1) \cdot (\bar{u}_L \frac{\lambda^a}{2} \gamma_\mu b_L)(\bar{b}_L \frac{\lambda^a}{2} \gamma_\nu u_L) , \\ \mathfrak{a} &= [\frac{\alpha_s(\mu^2)}{\alpha_s(p^2)}]^{1/b} , b = 11 - \frac{2}{3}n_f . \end{aligned} \quad (14)$$

The impact of WA on the continuous part of the spectrum is thus determined by the matrix elements of both colour singlet and colour octet operators:

$$\begin{aligned} \int_y^1 \frac{d\Gamma_{\text{annih}}(\eta)}{d\eta} d\eta &\simeq N \cdot \mathfrak{a}^{-4} f_B^2 \cdot (\frac{\mathfrak{a}}{\mathfrak{a}_y})^4 [g_{\text{singl}}(\mathfrak{a}_y^{9/2} - 1/9(\mathfrak{a}_y^{9/2} - 1)) - 2/3 g_{\text{oct}}(\mathfrak{a}_y^{9/2} - 1)] , \\ N &= \frac{G_F^2 |V_{ub}|^2 M_B^3}{4\pi} , \quad \mathfrak{a}_y = [\frac{\alpha_s(\mu_{\text{had}}^2)}{\alpha_s((1-y)m_b)^2}]^{1/b} , \quad \mathfrak{a} = [\frac{\alpha_s(\mu_{\text{had}}^2)}{\alpha_s(m_b^2)}]^{1/b} \end{aligned} \quad (15)$$

where we have used the notation

$$\begin{aligned} \langle B^- | \bar{u}_L \gamma_\mu b_L \cdot \bar{b}_L \gamma_\nu u_L | B^- \rangle &- \langle \bar{B}^0 | \bar{u}_L \gamma_\mu b_L \cdot \bar{b}_L \gamma_\nu u_L | \bar{B}^0 \rangle = \\ &= \mathfrak{a}^{-4} f_B^2 \cdot (v_{\text{singl}} p_{B_\mu} p_{B_\nu} - g_{\text{singl}} M_B^2 \delta_{\mu\nu}) \\ \langle B^- | \bar{u}_L \frac{\lambda^a}{2} \gamma_\mu b_L \cdot \bar{b}_L \frac{\lambda^a}{2} \gamma_\nu u_L | B^- \rangle &- \langle \bar{B}^0 | \bar{u}_L \frac{\lambda^a}{2} \gamma_\mu b_L \cdot \bar{b}_L \frac{\lambda^a}{2} \gamma_\nu u_L | \bar{B}^0 \rangle = \\ &= \mathfrak{a}^{-4} f_B^2 \cdot (v_{\text{oct}} p_{B_\mu} p_{B_\nu} - g_{\text{oct}} M_B^2 \delta_{\mu\nu}) . \end{aligned} \quad (16)$$

The quantities  $\mathfrak{a}^{-4}$  have been factored out to make the dimensionless matrix elements  $g$  and  $v$  constant in the heavy quark limit. The operators in eq.(16) are normalized at the low scale  $\mu_{\text{had}}$ . According to eq.(15) the total annihilation width and its part that lies in the endpoint region A are given by, respectively

$$\begin{aligned} \Delta\Gamma_{\text{annihil}} &\simeq \frac{G_F^2 |V_{ub}|^2 M_B^3}{4\pi} \cdot (g_{\text{singl}} - 0.4 g_{\text{oct}}) , \\ \Delta\Gamma_{\text{annihil}}|_A &\simeq \frac{G_F^2 |V_{ub}|^2 M_B^3}{4\pi} \cdot g_{\text{singl}} , \end{aligned}$$

The numerical factor in front of  $g_{\text{oct}}$  corresponds to the normalization point  $\alpha_s(\mu^2) = 1$  and  $\Lambda_{\text{QCD}} = 180 \text{ MeV}$ .

In Fig.4 we illustrate the spectrum of eq.(15) for the two typical ratios  $g_{\text{singl}}/g_{\text{oct}} = -1$  and  $g_{\text{singl}}/g_{\text{oct}} = 1/3$ ; if the matrix elements have opposite signs the overall effect

of annihilation is actually enhanced relative to the one in the endpoint region A. Finally we have added in Fig.4 a  $\delta$ -function at  $y = 1$  representing the two body decay  $B^\pm \rightarrow l\nu$ ; its finite height is designated to reflect the chirality suppression of this mode.

they had the same relative sign the overall effect of annihilation would be smaller than the one in the end point region A. Finally we have added in Fig.4 a  $\delta$ -function at  $y = 1$  representing the two body decay  $B^\pm \rightarrow l\nu$  (its finite height is designated to reflect the chirality suppression of this mode) and *ad hoc* some small nonleading perturbative corrections.

Turning to the discussion of real  $B$  mesons we have to concede that the numerical relation between the effects in the perturbative (B,C) and nonperturbative (A) regions depends on the concrete position of the borderline  $\mu_{\text{hadr}}$  between the two regimes. In other words we do not know reliably how close numerically we can approach the endpoint while still observing the perturbative  $dy/(1-y)$  behaviour.

In reality the scale  $m_c^2/m_b$  that determines the energy where semileptonic  $b \rightarrow c$  decays start to contribute is rather close to the hadronic scale  $\mu_{\text{hadr}}$  (see e.g. the discussion in ref. [9]). Therefore only a small interval is left above the charm threshold for the perturbative regime. This raises legitimate doubts whether our treatment in its present stage allows for a detailed quantitative analysis of the regions A and B *separately*.

## 4 Invariant Mass of the Hadronic Final State

Understanding the structure of the hadronic final state is obviously of great theoretical interest; in addition it provides important help in reducing the background in experimental studies<sup>6</sup>.

The expectation value for the invariant hadronic mass (squared) provides the simplest yardstick for the structure of the final state. There are perturbative contributions to it from gluon bremsstrahlung; they are infrared safe and thus calculable in a straightforward manner (see e.g. [10]). The situation becomes considerably more delicate for nonperturbative corrections. Although the *shape* of distributions – say the lepton energy spectra in semileptonic decays – generally gets  $1/m_b$  corrections due to the Fermi motion [9] of the  $b$  quark, *integrated* rates do not – provided they are expressed in terms of the *quark* rather than the *hadron* masses. To say it differently: while the leading corrections are indeed of order  $1/m_b$  they can be expressed completely in terms of  $\bar{\Lambda}$  – the asymptotic mass difference between the hadron and the heavy quark (for details see ref. [11]).

At this point we disagree with the conclusions of ref. [5]: whereas allowance was made there for  $1/m_b$  nonperturbative corrections to the total semileptonic widths, it was claimed that no such corrections arise in either the shape of the lepton spectrum or in the average hadronic invariant mass. The oversight in the latter claim is easily traced: the physical invariant mass in the decay is given by

$$M_{\text{hadr}}^2 = (p_B - q)^2 = M_B^2(2x - 1 + Q^2) \quad (17)$$

---

<sup>6</sup>We are grateful to Sheldon Stone for emphasizing this point to us.

where  $x$  and  $Q^2$  are as defined in Section 2, i.e. in terms of the mass and the four momentum of the beauty *hadron*. On the other hand, as shown in ref. [1] no  $1/m_b$  terms are present in the integrated distributions once they are expressed in terms of the mass and the four momentum of the beauty *quark*. Nontrivial corrections appear at the  $1/m_b^2$  level; their general expressions are given in ref. [12]. Therefore it is the quantity

$$\langle (p_b - q)^2 \rangle = \langle M_{\text{hadr}}^2 \rangle - 2\langle (p_b - q)(p_B - p_b) \rangle - \langle (p_B - p_b)^2 \rangle \quad (18)$$

rather than the invariant mass that is not subject to  $\mu_{\text{hadr}} m_b$  corrections; on the other hand the leading  $1/m_b$  corrections to the invariant mass squared of the hadrons in the final state are given by

$$\begin{aligned} \langle M_{\text{hadr}}^2 \rangle &\simeq m_q^2 + \frac{z_m(x)}{z_0(x)} \bar{\Lambda} m_b \ , \\ z_m(x) &= \frac{7}{10} - \frac{5}{2}x + 16x^2 - 16x^3 + \frac{5}{2}x^4 - \frac{7}{10}x^5 - 6x^2(1+x) \log \frac{1}{x} \ , \\ z_0(x) &= 1 - 8x + 8x^3 - x^4 + 12x^2 \log \frac{1}{x} \ , \quad x = (m_q/m_b)^2 \ . \end{aligned} \quad (19)$$

The numerical factor  $z_m/z_0$  varies from .7 to 2 for  $m_q$  between 0 and  $m_b$ ; it gives just the average energy of the quark  $q$  in the free semileptonic decay, measured in terms of  $m_b/2$ . This resolves the apparent puzzle raised in ref. [5]: the absence of operators yielding  $1/m_Q$  corrections was interpreted there as implying that no  $1/m_Q$  contributions arise in the invariant hadronic mass of the final state – although very intuitive physical arguments suggested otherwise. Note also that eq.(19) satisfies the heavy quark limit when  $m_b - m_c \ll m_c$ : it yields the mass  $(m_c + \bar{\Lambda})^2 \simeq M_D^2$ .

We will study here in particular how  $\langle M_{\text{hadr}}^2 \rangle$  differs in  $B_d$  and  $B^-$  decays. The effects of perturbative gluon bremsstrahlung in the quasifree amplitude cancel out in such a difference, hence the nonperturbative contributions are well defined here.

The contributions of WA to the total width scale like  $1/m_b^3$ ; it would then be natural to conjecture that the shift in  $M_{\text{hadr}}^2$  caused by WA is rather suppressed, namely of order  $\mu_{\text{hadr}}^3/m_b$  or even  $\mu_{\text{hadr}}^4/m_b^2$ . This would be in agreement with the observation that  $\langle M_{\text{hadr}}^2 \rangle$  is not or hardly modified by  $B \rightarrow l \nu_l$  and other decays into final states with low hadronic energies as described by the matrix elements of eq.(6), because  $M_{\text{hadr}}^2$  is very small in such transitions. On the other hand gluon emission changes the kinematics completely. Strictly speaking colour conservation requires the presence of at least two gluons for WA driven semileptonic decays. For the sake of simplicity we again ignore the colour indices and will discuss the general effect for the case of one gluon in the intermediate state.

Two observations would seem to lead to a startling conclusion: the processes with a real gluon, Fig.1b, and the one containing an on shell  $u\bar{u}$  pair produced via a virtual gluon, Figs.1c – when considered separately – induce large corrections to the decay probability  $\propto 1/m_b$ , see e.g. ref. [4]. Whereas the hadronic mass for the real gluon annihilation of Fig.1b is literally zero, the typical invariant mass square for two other cuts is  $m_{\text{sp}} \cdot m_b$  because the main contribution comes from  $\vec{k}_u \sim m_b$ . Yet the latter

produce a (negative) correction to the width of order  $1/m_b$ ! This would imply that through gluon emission WA can induce a negative correction to  $M_{\text{hadr}}^2$  as large as  $\alpha_s \cdot \mu_{\text{hadr}}^2$  – in clear conflict with the conjecture stated above.

We will show now that this line of reasoning based on a naive physical picture of the decay process is fallacious. We will also explain what the loophole is in the argument and that consistent calculations yield the correct result even in such a simple ansatz.

The corrections to  $\langle M_{\text{hadr}}^2 \rangle$  that is generated by the diagram in Fig.2 are given by multiplying the expression for the diagram with  $M_{\text{hadr}}^2 = (p_B - q)^2 = M_B^2(2x - 1 + Q^2)$ . This factor is analytical; therefore the integral of  $W(2x - 1 + Q^2)$  over  $x$  can be evaluated via the usual manipulations like deforming the integration path to the contour  $\beta$  in Fig.3. In fact this factor merely cancels out the pole associated with the gluon propagator. Again the result has a regular limit when the spectator momenta vanish. In this limit one arrives at the integration of a function that contains a double pole at  $x = (1 - Q^2)/2$ . The analysis then proceeds exactly like in Section 2 and yields a regular four fermion amplitude similar to the original one. Its matrix element then scales like  $1/m_b^3$  and therefore it leads to corrections to  $\langle M_{\text{hadr}}^2 \rangle$  only of order  $\mu_{\text{hadr}}^3/m_b$  – as conjectured. The essential difference as compared to the total probability is that the integration over  $Q^2$  has now the form  $dQ^2(1 - Q^2)$ : the main contribution thus comes from high momenta hadronic states. The correction is therefore given by the matrix element of an operator different from the one that expresses the WA contribution to the total width.

We have thus shown that WA does *not* induce corrections of order  $\mu_{\text{hadr}}^2$  to  $M_{\text{hadr}}^2$  after all: they are suppressed at least by one extra power of  $1/m_b$ . This result leads immediately to the question of what could have gone wrong with the simple argument sketched above that was based on a nonrelativistic picture of the  $B$  meson?

The short answer is that the contributions from the two cuts across the  $u\bar{u}$  pair in the quark diagram become more singular in the nonrelativistic limit and that they have both signs. Consider for example this forward scattering amplitude for  $b\bar{u} \rightarrow b\bar{u}$  with the momentum of the spectator  $\bar{u}$  being identical in the initial and final state: the left hand and the right hand cut taken separately both become infinite. This singularity is – most conveniently and naturally – regulated by assigning slightly different momenta to the initial and final antiquarks; both cut contributions are then finite though by themselves enhanced by the small difference in the spacelike momenta of the initial and final  $\bar{u}$  that appears in the denominator. It is easy to see that this enhancement which reflects the original singularity disappears from the sum of the two terms when one calculates the total decay probability. Such a cancellation has to be expected on general grounds: for the singularity reflects the emergence of a strong phase due to one-gluon rescattering which does not affect the width directly. On the other hand the singularity becomes relevant in yielding a finite contribution to the invariant mass! For the difference in the spectator momenta produces a corresponding difference in the invariant mass for the two cuts; its small size is compensated for by the nonrelativistic enhancements of the two separate cut contributions. Thus we obtain a contribution to the invariant mass even in the limit of identical velocities of the initial and final  $\bar{u}$  antiquark that is finite and positive and thus compensates for

the negative contribution obtained above!

It is instructive to analyze this problem from a more general perspective as well because the apparent paradox can be formulated in very general terms without referring to a nonrelativistic description. The probability for the decay of a  $b$  quark into a final state containing an on-shell gluon without a  $u\bar{u}$  pair is of course positive and scales like  $1/m_b$ . As the WA correction to the total decay width is of order  $1/m_b^3$ , the probability for  $b$  quark decays to final states with a  $u\bar{u}$  pair has to scale like  $1/m_b$  and be negative. The typical mass square of such states is of order  $\mu_{\text{hadr}} m_b$ ; therefore they lead to a large negative correction to  $\langle M_{\text{hadr}}^2 \rangle$ . This paradox has an interesting solution that can be formulated in obvious analogy to problems in ordinary quantum mechanics when one considers for simplicity a system with a discrete spectrum of states rather than a continuous one. For calculating the corrections to  $\langle M_{\text{hadr}}^2 \rangle$  one has to consider not only the corrections to the probability of the decays to  $u\bar{u}$  states but also to include the shift in their energy levels. Because the massless gluon has an energy less than the energy of the  $q\bar{q}$  pair, its exchange *raises* the invariant mass of the latter, and therefore provides the missing positive contribution to  $M_{\text{hadr}}^2$ .

This standard quantum mechanical effect is of course automatically contained in a computation based on Feynman diagrams. What we have demonstrated above is that the necessary cancellations arise also as long as one deals carefully with the singularities that emerge in a nonrelativistic treatment of quark diagrams. There is a general lesson to be learnt from it: *a priori* one has to be very careful in drawing conclusions about small nonperturbative corrections appearing on top of large free parton amplitudes. More specifically our analysis in this section shows once again that some effects of WA can emerge in rather delicate ways. A failure to treat properly all possible contributions where identical  $q\bar{q}$  pairs are present can easily lead to a dramatic overestimate of the impact of WA. To cite but one example: WA can certainly induce scattering phases in some exclusive modes and thus affect their transition rates significantly; yet from such an observation alone one cannot infer that WA is numerically important in inclusive  $B$  and  $D$  decays.

To conclude this section we add the comment that an analysis similar to the one outlined here can be applied also to interference effects in semileptonic decays of heavy flavour baryons. PI contributes first on the level  $1/m_b^3$  and thus can cause a shift in  $M_{\text{hadr}}^2$  at most of order  $\mu_{\text{hadr}}^3/m_b$ .

## 5 Summary and Outlook

When the first data on  $D$  meson lifetimes showed a large difference between  $\tau(D^+)$  and  $\tau(D^0)$  it was immediately realized that WA can shorten  $\tau(D^0)$  considerably – if its inherent helicity suppression can be overcome. In particular it was suggested that the presence of gluons could enhance WA contributions. Ref. [8] considered the perturbative emission of energetic gluons. Yet our analysis in ref. [4] showed that perturbative gluon radiation can vitiate helicity suppression only at a small numerical level of order  $\alpha_S(m_b^2)$ .

Another attempt [13] invoked nonspectator contributions in the form of soft gluons in the wavefunction of heavy flavour hadrons. In the present paper we have discussed



in considerable detail the impact of weak annihilation on weak decays of heavy flavour hadrons with the main emphasis on semileptonic  $b \rightarrow u$  decays, since those are easier to treat theoretically. The results we found are in broad qualitative agreement with the ansatz of ref. [13] – as long as it makes sense at all to speculate about nonperturbative dynamics in terms of individual gluons. Gluons in the wavefunctions could – in an oversimplified picture – generate the nontrivial intermediate states saturating the local four fermion operator in eq.(6): for the quark current then annihilates only the constituent quarks in the wavefunction leaving extra quanta in the final state. Of course the agreement between our analysis and that of ref. [13] is only of a very rough qualitative nature. Among other differences we have shown that the relative weight of such nonspectator contributions scales like  $1/m_Q^3$  and not like  $1/m_Q$  or  $1/m_Q^2$  as conjectured in ref. [13].

More specifically we have shown that nonperturbative effects could generate a significant difference in  $\Gamma(B^0 \rightarrow l\nu + X_u)$  vs.  $\Gamma(B^- \rightarrow l\nu + X_u)$  that is *not* suppressed by the mass of the lepton and *a priori* could be 6 times as large (for  $g \simeq 1/3$ ) as  $\Gamma(B \rightarrow \tau\nu_\tau)$  for both electron and muon decays; this would represent a significant contribution, namely  $\sim 20\%$  of the total  $b \rightarrow ul\nu$  width. Furthermore this difference is concentrated in a narrow slice – with a width  $\mu_{\text{hadr}}/2$  – just below the endpoint of the lepton energy spectrum; for real B decays it is the region starting around 2.3 GeV. Therefore if one studies only the region *above* the  $b \rightarrow c$  endpoint the relative weight of WA in semileptonic  $B^-$  decays gets enhanced and could actually become large. This additional contribution in  $B^\pm$  decays can significantly affect the value of  $|V_{ub}|$  that up to now has been derived from semileptonic spectra averaged over charged and neutral  $B$ 's.

The exact numerical size of the WA contributions to the lepton spectrum depends on the magnitude of two matrix elements; one involves colour singlet and the other colour octet operators. In the endpoint region, the domain of nonperturbative dynamics, the colour singlet operator dominates. There exists also a “logarithmic” tail extending below the endpoint region where a perturbative treatment applies; colour octet operators play a larger role there. It should be noted that numerically annihilation effects could be larger than the flavour independent effects identified in ref. [9] although the latter are formally leading in  $1/m_b$  (see ref. [6]).

Obviously inclusive charm decays can be analyzed in a quite analogous fashion since the heavy quark symmetry ensures that the same four fermion matrix elements enter in the decay rates of charm hadrons; the only difference arises due to hybrid renormalization which mixes the colour singlet and octet operators and which can be calculated perturbatively. Of course the nonperturbative corrections are much larger in  $D^+ \rightarrow l\nu X$  vs.  $D^0 \rightarrow l\nu X$  than in  $B^- \rightarrow l\nu X$  vs.  $B^0 \rightarrow l\nu X$  decays due to  $1/m_c > 1/m_b$  and also  $|V(cd)/V(cs)| > |V(ub)/V(cb)|$ ; even more prominent effects are expected in  $D_s \rightarrow l\nu X$  decays. This makes such an analysis of charm decays more challenging as well as more intriguing. We will report on it in a future paper [7].

Our results are rather qualitative or at best semi-quantitative in nature. This is not surprising since we are dealing with nonperturbative corrections in general and the size of nonfactorizable matrix elements in particular. At present we have little theoretical guidance in evaluating these features in a numerically reliable fashion:

folklore has it that the nonfactorizable parts in matrix elements are suppressed by the number of colours; thus one expects  $g_{\text{singl}}$  to be less than unity in the real world. At the same time however the suppression of  $g_{\text{oct}}$  may appear to be softer. On the other hand QCD sum rules tend to suggest that nonfactorizable contributions to the  $\delta_{\mu\nu}$  term are in general smallish in heavy flavour mesons [14]; yet further studies are obviously required before reliable conclusions can be drawn.

Semileptonic  $\Lambda_b$  decays provide an intriguing lab to study effects due to both WS and PI. They can be analyzed in close analogy to the discussion given in this paper. For example PI affects (KM suppressed) semileptonic  $\Lambda_b$  decays and it does so mainly in the endpoint region – as it is the case for WA in  $B$  meson decays. Since WS is *not* generally helicity suppressed its relative impact on preasymptotic corrections could be sizeable. Yet for that very reason one cannot expect nonperturbative low energy scale physics to play a prominent and easily identifiable role in  $\Lambda_b$  decays. Furthermore here we do not have an unambiguous yardstick as we had it in the comparison between charged and neutral  $B$  mesons. All of this will make it considerably more difficult to separate out various preasymptotic effects.

$B_c$  mesons constitute a very intriguing dynamical system. With the charm quark mass providing an infrared cut-off  $B_c$  decays can be treated in perturbation theory. WA is not KM suppressed here and the hadronic matrix element that determines the strength of WA can be calculated both in its factorizable and nonfactorizable part. Furthermore final states from WA that contain a  $c\bar{c}$  pair should be experimentally separable from those with a gluon (or gluons) instead in the final state. For the latter one can then invoke the nonrelativistic calculation of ref. [8] with its enhancement factor  $(m_b/m_c)^2$ . These calculations are not trivial, but they appear to be doable in a straightforward way. Obtaining a sample of  $B_c$  decays sufficiently large to make the studies referred to above feasible represents a stiff experimental challenge, yet it might not be beyond our reach.

There exist many cross references between a theoretical understanding of semileptonic and nonleptonic heavy flavour decays. More specifically the preasymptotic corrections that we have been discussing in this note appear already in the KM allowed  $b \rightarrow c$  transitions, namely WA in  $B_d$  and  $B_s$ , PI in  $B^\pm$  decays. We then infer that the impact of WA and PI on total nonleptonic decay rates can be expressed in terms of matrix elements of two four fermion operators that differ solely in the way colour flows through them. We also conjecture that the weight of WA in the total width is determined mainly by the same nonfactorizable piece coming from nonperturbative low energy dynamics that can be extracted from a detailed study of semileptonic  $B$  decays. A considerably larger correction in nonleptonic decays is expected from PI which appears already in factorizable contributions (see refs. [3, 4] for details). For this reason we view our earlier estimate as reliable – unless an unexpectedly large difference in the lepton spectra in  $B^0$  vs.  $B^-$  decays forces us to revise the common wisdom about the numerical validity of the factorization approximation in the real world; in that case one had to reconsider the hierarchy of the preasymptotic corrections outlined above.

**ACKNOWLEDGEMENTS:** We gratefully acknowledges illuminating dis-

cussions and exchanges on the subject of this paper with several colleagues, and in particular with V.Braun, Yu.Dokshitzer, A.Vainshtein and A.Johansen. This work was supported in part by the National Science Foundation under grant number PHY 92-13313.

## References

- [1] I.I.Bigi, N.G.Uraltsev, A.Vainshtein, *Phys.Lett.* **293B** (1992) 430;  
B.Blok, M.Shifman, *Nucl.Phys.***B399** (1993)441, 459;  
I.I.Bigi, B.Blok, M.Shifman, N.G.Uraltsev, A.Vainshtein, **preprint** UND-HEP-92-BIG07, TPI-MINN-92/67-T.
- [2] M.A.Shifman, M.B.Voloshin, *Sov.Journ.Nucl.Phys.* **41** (1985) 120.
- [3] M.A.Shifman, M.B.Voloshin, *Sov.Phys.ZhETF* **64** (1986) 698; *Sov.Journ.Nucl.Phys.* **45** (1987) 292.
- [4] I.I.Bigi, N.G.Uraltsev, *Phys.Lett.* **280B** (1992) 120; more details can actually be found in the preprint version, UND-HEP-92-BIG02.
- [5] J.Chay, H.Georgi and B.Grinstein, *Phys.Lett.* **B247** (1990) 399.
- [6] I.I.Bigi, N.G.Uraltsev, A.Vainshtein, *Phys.Lett.* **293B** (1992) 430.
- [7] I.I.Bigi, N.G.Uraltsev, **preprint** UNDHEP-BIG03 (1993).
- [8] M.Bander,D.Silverman, A.Soni, *Phys.Rev.Lett.* **44** (1980) 7, 962(E).
- [9] I.I.Bigi, M.Shifman, N.G.Uraltsev, A.Vainshtein, *Phys.Rev.Lett.* **71** (1993) 496.
- [10] G.Altarelli et al., *Nucl.Phys.* **B208** (1982) 365.
- [11] I.I.Bigi, M.Shifman, N.G.Uraltsev, A.Vainshtein, *in preparation*.
- [12] B.Blok, L.Koyrakh, M.Shifman, A.Vainshtein, **preprint** TPI-MINN-93/33-T.
- [13] H.Fritzsch, P.Minkowski, *Phys.Lett.* **90B** (1980) 455.
- [14] V.Braun, private communication.

## Figure Captions

**Fig.1** Parton diagrams for the simplest WA processes:

- a) “Parton annihilation” at tree level.
- b) Conventional gluon annihilation.
- c) Interference in the “quasifree” mode  $b \rightarrow u + \bar{q}q' (l\nu_l)$  induced by the gluon WA

amplitude.

**Fig.2** Lowest order perturbative diagram for the WA-induced correction to correlators  $W$  in eq.(1).

**Fig.3** The complex  $x$  plane. Crosses denote the poles representing singularities of the one gluon annihilation amplitude. The heavy line shows the cut of the Jacobian in eq.(2).

**Fig.4** Shape of the difference in the lepton spectra for  $B^\pm \rightarrow X_u l \nu$  vs.  $B^0 \rightarrow X_u l \nu$  decays corresponding to  $g_{\text{oct}}/g_{\text{singl}} = -1/2$  (solid line) and  $g_{\text{oct}}/g_{\text{singl}} = 1/3$  (dashed line). Below region A ( $E_1 < 2.6 \text{ GeV}$ ) the curves describe the spectrum given by eq.(15) with  $\Lambda_{\text{QCD}} = 180 \text{ MeV}$ . The thick vertical line at  $E_l = M_B/2$  represents the monochromatic line from the chirality suppressed two body mode  $B^\pm \rightarrow l \nu$ .

This figure "fig1-1.png" is available in "png" format from:

<http://arXiv.org/ps/hep-ph/9310285v1>

This figure "fig1-2.png" is available in "png" format from:

<http://arXiv.org/ps/hep-ph/9310285v1>

This figure "fig1-3.png" is available in "png" format from:

<http://arXiv.org/ps/hep-ph/9310285v1>

This figure "fig1-4.png" is available in "png" format from:

<http://arXiv.org/ps/hep-ph/9310285v1>